Common Product Self-Organizing Map – CPSOM

Luciano da Fontoura Costa
luciano@ifsc.usp.br

São Carlos Institute of Physics – DFCM/USP

4th Nov 2021

Abstract

Self-organizing maps (SOMs) are non-supervised neuronal networks based on several biological principles. This type of network has been found to provide good performance for several pattern recognition networks, including sound and image applications. Based on multiset concepts, the common product and respective correlation has been found to typically outperform extensively employed methods such as the cosine similarity, inner producte and cross-correlation. In the present work, we bring these two research areas together by developing a self-organized network whose neurons incorporates the common product instead of the inner product as a model of synaptic transduction and integration. It is shown that this type of obtained neuronal networks present several interesting characteristics including partitioned regions with more uniform areas as well as enhanced cluster separation.

"*Soap bubbles, those masters of space.*'

LdaFC

1 Introduction

Though it often goes not realized, pattern recognition (e.g. [1, 2, 3]) underlies a good deal of human (as well as other living beings) activities which include but are by no means limited to: reading these lines, hearing a distant bird song, and smelling springtime. Virtually every of these perceptions (e.g. [4, 5]) are achieved through intricate neuronal systems perfect along millions of years.

Though the ubiquity of patter recognition may not be so often acknowledge, its importance is nonetheless strongly reflected on the continuing and intense efforts aimed at achieving effective machine learning correlates. Though dating back to William James's studies in the the XIX century, the important area of neuronal pattern recognition underwent important developments more recently in the *deep learning initiative* (e.g.[6, 7]).

One of the neuronal models that have played an important role in patter recognition, including respectively to deep learning, is the Kohonen's self-organizing map approach (e.g. [8]), which was strongly motivated by findings from neurobiology and neuroscience (e.g. [5, 4], including not only at the neuronal cell level, but also in the way in which these cells organize themselves topographically along the mammals cortex so that nearby neurons tend to have similar operation, and in the ultimate consequences of resulting neuronal modules associated to specific patterns, which accounts for the *self-organized map* — SOM — name.

In their simplest form, these neuronal networks are characterized by having its input layer connected directly to single 'cortical' layer, where linear neurons are distributed uniformly. Each neuron performs an inner product between its current weights and the input pattern. Therefore, the neuronal output can be understood to reflect the similarity between the input pattern and the current weight values.

The learning procedure goes as follows. For each presented stimulus, the weights modification takes place around the cell presenting the highest response. These modifications are often performed through a simple rule taking into account the vector diference between the input pattern and the weights, while also taking into account a learning rate α and the distance between the neighboring cells and that which presented the highest output. The further away a cell is from the winning cell, the less intense the weight correction is performed. The repetition of this learning stage, in presence of possible relaxation of the learning parameters, has been verified to yield partitions of the cortical layer into portions corresponding to the types of presented patterns.

Though the inner product has been extensively adopted in neuronal networks as the means to quantify the similarity between each pattern and the weight distribution, there is a virtually infinite number of other possible similarity approaches. More recently [9, 10, 11], the classic Jaccard index [12], whose application was largely limited to binary or categorical data, was adapted to take into account real, possibly negative values, as allowed by the consideration of multiset concepts and methods (e.g. [13, 14, 15, 16, 17, 18]) . Since this index has been shown not to take into account the relative interiority between two sets [9], a respective enhancement was proposed [12] that consists of multiplying the real-valued Jaccard index by the interiority index, giving rise to the *coincidence index* [9, 10, 11]. The application of these indices to traditional important problems such as template matching pattern recognition [19], complex network representation and modularity enhancement [20], as well as clustering [21], has indicated that the aforementioned multiset-based similarity approaches have enhancement for pattern recognition thanks to their increased sensitivity and accuracy underlain by a respective more strict quantification of pairwise similarity between virtually any mathematical structure, including sets, multisets, vectors, functions and fields.

In addition, complementary studies [22] indicated that the multiset-based similarity indices have enhanced performance also when the traditional inner product traditionally adopted as the means for modeling and performing synaptic transduction and integration. In particular, individual artificial neuronal cells employing the multiset-based approaches, and especially the coincidence index-based common product, have substantially improved capabilities for achieving accurate, strict and specific recognition of image patterns

In the present work we take a step further and investigate the performance of the multiset-based approaches respectively to whole ensembles of neurons, or neuronal networks, while focusing on the SOM approach.

We start by briefly presenting the multiset-based similarity indices, and then proceed to illustrating their enhanced potential when applied to a SOM network. The results obtained so far indicate that, when compared to the traditional inner used for product learning rule and recognition, the multiset-based approaches tend to yield more balanced partitioning of the cortical space, as well as increasingly separated frontiers between the resulting modules associated to the input patterns. These abilities are understood to be a direct consequence of the enhanced capability of the multiset-based similarity measurements, including the common product, respectively to more strict and accurate similarity quantification.

2 The Common Product and Correlation

We start with the *classic cross-correlation* between two functions $f(x)$ and $g(x)$:

$$Corr(f,g)\,[y] = \int_{-\infty}^{\infty} f(x)g(x-y)dx \tag{1}$$

The *basic Jaccard index* between two sets with non-negative multiplicity can be defined as:

$$\mathcal{J}(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

where A and B are any two sets to be compared. It can be verified that $0 \leq \mathcal{J}(A,B) \leq 1$.

The *homogeneity* or *interiority index* can be expressed as:

$$\mathcal{I}(A,B) = \frac{|A \cap B|}{\min\{|A|,|B|\}} \tag{3}$$

with $0 \leq \mathcal{I}(A,B) \leq 1$ and $0 \leq \mathcal{J}(A,B) \leq \mathcal{H}(A,B) \leq 1$.

The combination of the basic Jaccard and the interiority indices yields the *coincidence index*, proposed in [9]:

$$\mathcal{C}(A,B) = \mathcal{J}(A,B)\,\mathcal{I}(A,B) \tag{4}$$

The Jaccard index extended to multisets with non-negative multiplicities can be expressed as:

$$\mathcal{J}_M(A,B) = \frac{\sum_{i=1}^{N} \min(a_i, b_i)}{\sum_{i=1}^{N} \max(a_i, b_i)} \tag{5}$$

with $0 \leq \mathcal{J}_M(A,B) \leq$. Therefore, this index is capable of taking into account the multiplicity of the elements in the involved multisets.

The extension of the Jaccard index to real, possibly negative-valued multiplicities, involves the following developments [9, 10].

In particular, the possibility to have positive and/or negative multiplicity values now requires the application of a binary operator analogous to the inner product in function spaces, in which the multiplicities are properly mirrored among the four quadrants depending on their signs [9, 10]. This operator, which will be referenced here as *common product*, can be expressed as:

$$\ll f(x), g(x) \gg = \int_{-\infty}^{\infty} s_f s_g \min(s_f f(x), s_g g(x))dx \tag{6}$$

This binary operator (in the sense of taking two arguments) corresponds to the direct counterpart of inner product between multisets with non-negative multiplicities.

We shall also adopt the following operator:

$$f(x) \circ g(x) = \int_{-\infty}^{\infty} \max(s_f f(x), s_g g(x))dx \tag{7}$$

From the above results, the *multifunction convolution* [9, 10] (*mconvolution*) of two functions can now be derived:

$$f(x) \blacksquare g(x)\,[y] = \int_{-\infty}^{\infty} \ll f(x) g(y-x) \gg dx \tag{8}$$

Similarly, we derive the *multifunction correlation*, or *mcorrelation* as:

$$f(x) \square g(x)\,[y] = \int_{-\infty}^{\infty} \ll f(x) g(x-y) \gg dx \tag{9}$$

It is also possible [10] to define:

$$f(x) \circledast g(x)\,[y] = \int_{-\infty}^{\infty} \left[f(x) \circ g(x-y) \right] dx \tag{10}$$

We also have:

$$A_f = \int_{-\infty}^{\infty} |f(x)| dx \tag{11}$$

$$A_g = \int_{-\infty}^{\infty} |g(x)| dx \tag{12}$$

Now, the multifunction correlation, when normalized by the above function, yields:

$$f(x) \square g(x)\,[y] = \frac{\int_{-\infty}^{\infty} \ll f(x), g(x-y) \gg dx}{\int_{-\infty}^{\infty} f(x) \circledast g(x-y) dx} \tag{13}$$

observe that the integrand corresponds to the generalized Jaccard index adapted to cope also with negative multiplicities, here called *negative multiset Jaccard index*.

The interiority index also needs to be adapted to possibly negative multiplicities. First, we make:

$$\ll f(x), g(x) \gg_{+} = \int_{S_+} \min(s_f f(x), s_g g(x)) dx \tag{14}$$

where $S_+ = \{x | s_f(x) s_g(x) > 0\}$, and then we can write:

$$\mathcal{I}(f(x), g(x)) = \frac{\ll f(x), g(x) \gg_{+}}{\min\{A_f, A_g\}} \tag{15}$$

The respective combination with the interiority index yields the *coincidence index* [9, 10], written as:

$$\mathcal{C}_N(f(x), g(x)) = \mathcal{J}_N(f(x), g(x))\, \mathcal{I}(f(x), g(x)) \tag{16}$$

Now, when performing template matching between two functions, it may be interesting to consider the interiority also when $s_f(x) s_g(x) < 0$. In this case, it is possible to take the absolute value of the object function as argument of the interiority index above, which is the case considered henceforth in this work.

Each of these indices lead to respective multifunction convolutions and correlations, which involve sliding one function with respect to the other while calculating the respective index, followed by the respective integration. Actually, many neuronal networks including the SOM and other deep learning architectures, can be understood as being based on the convolution or correlation between the stimuli and the cells synaptic weights.

3 Common Product-Based Self-Organizing Maps – CPSOM

The common product-based self-organizing map — CPSOM — model can be understood as a SOM in which the inner product has been replaced by some common product-based similarity index, including the real-valued Jaccard and coincidence indices. All other aspects, including stimuli presentation and training, are shared with the traditional SOM approach.

In this section we consider the real-valued Jaccard and coincidence indices, as well as the classic inner product-based cross correlation method.

The input consisted of guassian paterns placed at three distinct positions in the input space (or retina), as illustrated in Figure 1. Both the input ('retina') and neuronal ('cortex') layers have dimension 20 × 20.

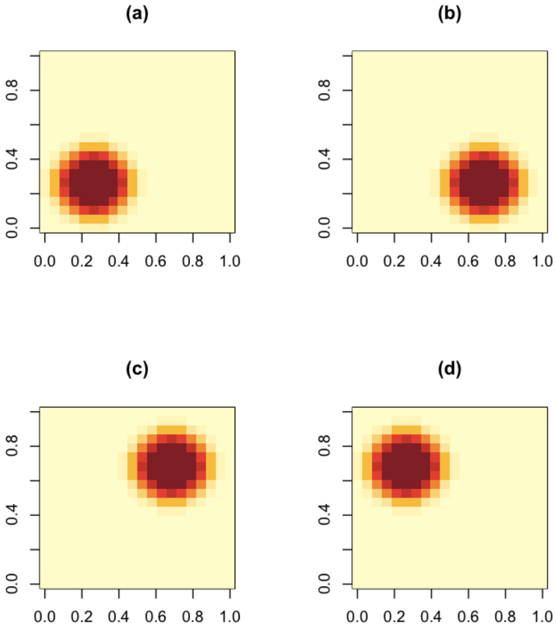

Figure 1: The four types of patterns considered for training the CPSOM were obtained from these four gaussians by addition of symmetric uniform noise and small spatial displacements of each pattern.

Several instances of these three types of patterns were then obtained by adding symmetric noise and positional displacements of each stimuli before respective presentation to the CPSOM architecture. The following results were obtained for 100 presentations, with the following learning rule:

$$W^{[t+1]} = W^{[t]} + \left[X - W^{[t]} \right] \; \alpha \; e^{-\beta d} \tag{17}$$

where d is the distance from the current cell to the respective winning cell at a given epoch t. We have adopted $\alpha = 0.2$ and $\beta = 0.1$. The neighborhood considered around each winning cell consists of those neurons that are no further apart than $R = 10$.

After training, each cell was associated to the noiseless version of the pattern to which it produces the highest response, therefore defining the partition of the cortex space. We also considered an additional result to be visualized corresponding to the intensity of the maximum response of each of the cells to the associated noiseless pattern.

Figures 2, 2, and 2 illustrates the obtained partitioned cortex as well as the respective intensity plots respectively to the adoption of the real-valued Jaccard, coincidence and classic correlation similarities

The result obtained for the coincidence method presents partitioned regions not only occupying more similar areas, but also better separated in the intensity plot. The second best approach was that based on the real-valued Jaccard index, followed by that adopting the classic cross correlation between the weights and stimuli.

In addition to these possibly enhanced features regarding the resulting properties of the partitioned cortex, it would also be reasonable to expect that subsequent pattern recognition will be performed in a more strict and accurate manner by the coincidence-based CPSOM, given that this index has shown enhanced capabilities when applied to other pattern recognition related problems [19, 21, 20, 22].

FIgure 5 presents the resulting cortex and intensity plot obtained by using the coincidence-based CPSOM, but with the minimum part of the Jaccard sub-index being taken to the second power, as suggested in [9]. An even better separation of the clusters can be observed.

4 Concluding Remarks

Effective, generic pattern recognition constitutes one of the greatest challenges that motivates much of current science and technology, given the ubiquity and importance of this ability for many human activities. In particular, significant recent results in the area of deep learning have allowed several difficult problems to be effectively tackled and solved. A good deal of these advances have been motivated and relied on concepts and advances from the closely related area of neuronal networks (e.g. [23]).

Oftentimes, the synaptic transduction and integration by individual neuronal cells have been modeled and/or implemented artificially in terms of the classic inner product. In the present work, we set out at investigating the potential of alternative mutiset-based measurements of similarity as substitutes for the inner product in the important neuronal network known as SOM.

We implemented a relatively small CPSOM based on the multiset-derived real-valued Jaccard and coincidence indices, as well as on the classic inner product-based classic cross correlation. Four types of patterns, subject to respective modifications by symmetric uniform noise as well as by small positional displacements, were then fed into each of the considered CPSOMs architectures. The results, evaluated visually from the obtained cortical partitions and intensity plots, suggest that the multiset methods, especially that based on the coincidence index, tend to yield more balanced partitions, in the sense of the respective regions presenting more similar areas. In addition, more markedly separations can ben obtained in the intensity plot obtained by using the coincidence index method. These results concern the learning stage, but Enhanced performance could also be expected when applying the obtained maps to recognize new patters.

Though our discussion has been based on a specific experiment reported in Section 3, more experiments (not shown) have been performed that tend to corroborate the observed advantages of the coincidence index-based CPSOM. Additional studies are being conducted incorporating more objective performance quantifications. other types of stimuli, as well as cortex sizes, and results should be communicated opportunely.

Acknowledgments.

Luciano da F. Costa thanks CNPq (grant no. 307085/2018-0) and FAPESP (grant 15/22308-2).

References

[1] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification.* Wiley Interscience, 2000.

[2] K. Koutrombas and S. Theodoridis. *Pattern Recognition.* Academic Press, 2008.

[3] L. da F.Costa. *Shape Classification and Analysis: Theory and Practice.* CRC Press, Boca Raton, 2nd edition, 2009.

[4] E. B. Goldstein. *Sensation and Perception.* Cengage Learning, 2012.

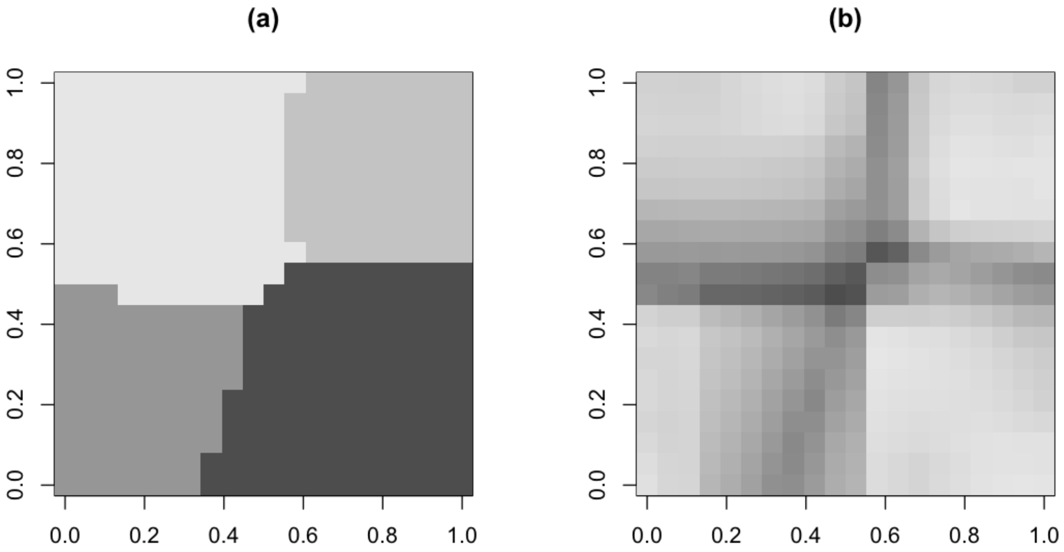

Figure 2: The real-valued Jaccard-basedCPSOM 'cortex' partitioned into the four types of patterns (a), as well as the respectively maximum intensities (b) obtained for the associated patterns. A good separation of the pattern types has been obtained, though being a little blurred.

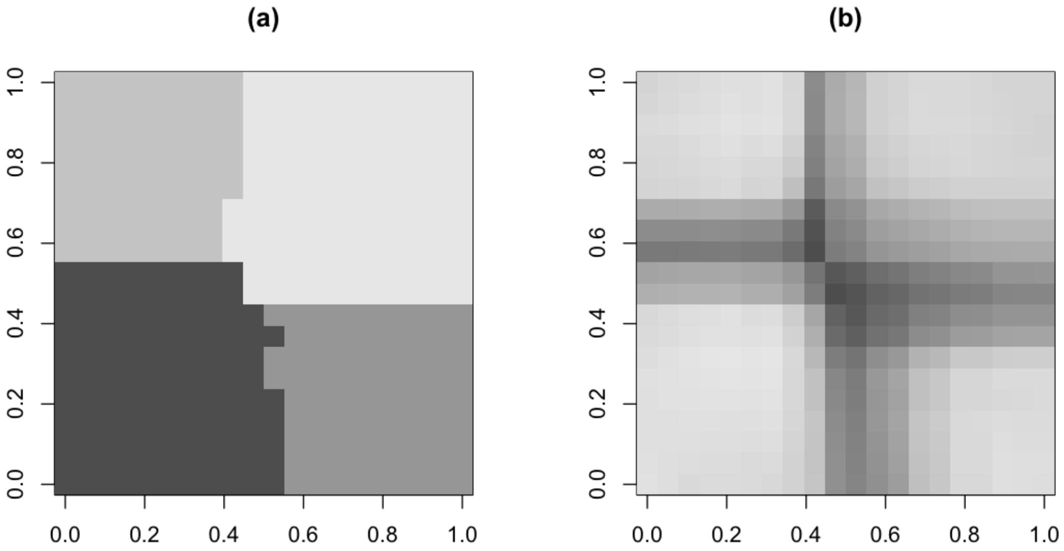

Figure 3: The coincidence-based CPSOM 'cortex' partitioned into the four types of patterns (a), as well as the respectively maximum intensities (b) obtained for the associated patterns. Not only the four regions present markedly similar areas, but also the interstices between then in the respective intensity plot are wider, indicating better separation of the four types of patterns.

[5] E. R. Kandel, J. H. Schwartz, T. M. Jessel, S. A. Siegelbaum, and A. J. Hundspeth. *Principles of Neural Science*. McGraw-Hill, 2009.

[6] J. Schmidhuber. Deep learning in neural networks:an overview. *Neural networks*, 61:85–117, 2015.

[7] G. E. Hinton. Training products of experts by minimizing contrastive divergence. *Neural computation*, 14(8):1771–1800, 2002.

[8] T. Kohonen. *Self-Organization and Associative Memory*. Springer Verlag, 2009.

[9] L. da F. Costa. Further generalizations of the Jaccard index. https://www.researchgate.net/publication/355381945_Further_Generalizations_of_the_Jaccard_Index, 2021. [Online; accessed 21-Aug-2021].

[10] L. da F. Costa. Multisets. https://www.researchgate.net/publication/355437006_

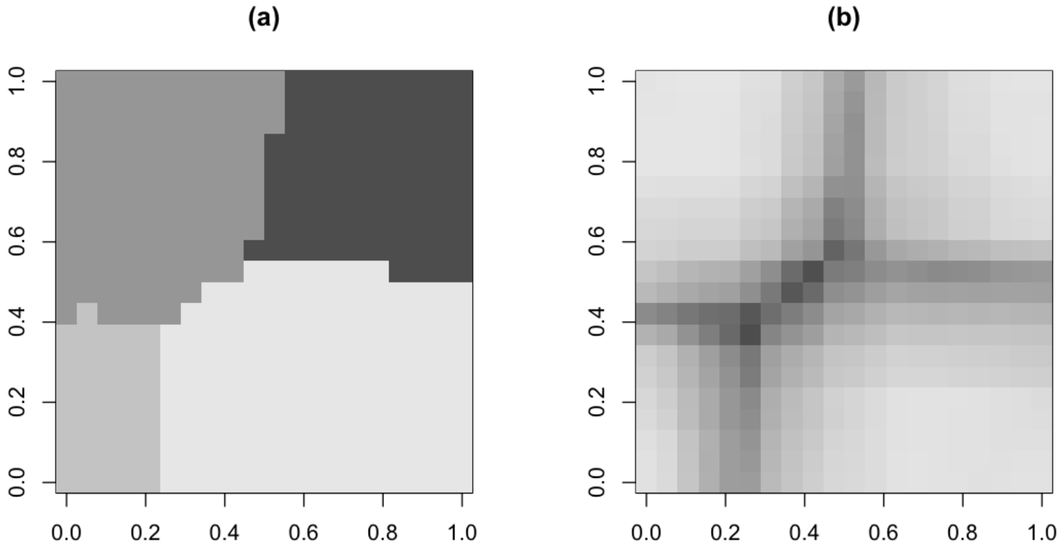

Figure 4: The inner product-based CPSOM 'cortex' partitioned into the four types of patterns (a), as well as the respectively maximum intensities (b) obtained for the associated patterns. Not only the obtained patter regions occupy markedly different areas, but also the pattern separations in the intensity plot are much weaker and blurred than obtained for the coincidence CPSOM or even the real-valued Jaccard CPSOM.

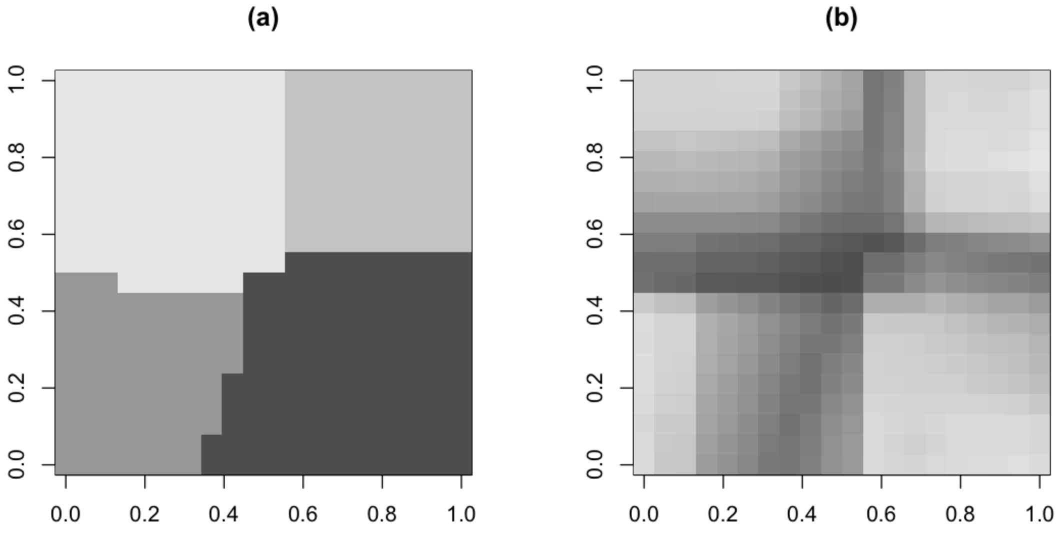

Figure 5: The partitioned cortex and respective intensity plot obtained by the coincidence-based CPSOM, but having the minimum operation (numerator) of the involved real-valued Jaccard squared. In addition to partitioned regions with similar areas, an even better separation of the clusters can be observed.

Multisets, 2021. [Online; accessed 21-Aug-2021].

[11] L. da F. Costa. On similarity. https://www.researchgate.net/publication/355792673_On_Similarity, 2021. [Online; accessed 21-Aug-2021].

[12] P. Jaccard. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin de la Société vaudoise des sciences naturelles*, 37:547–549, 1901.

[13] J. Hein. *Discrete Mathematics*. Jones & Bartlett Pub., 2003.

[14] D. E. Knuth. *The Art of Computing*. Addison Wesley, 1998.

[15] W. D. Blizard. Multiset theory. *Notre Dame Journal of Formal Logic*, 30:36—66, 1989.

[16] W. D. Blizard. The development of multiset theory. *Modern Logic*, 4:319–352, 1991.

[17] P. M. Mahalakshmi and P. Thangavelu. Properties of multisets. *International Journal of Innovative Technology and Exploring Engineering*, 8:1–4, 2019.

[18] D. Singh, M. Ibrahim, T. Yohana, and J. N. Singh. Complementation in multiset theory. *International Mathematical Forum*, 38:1877–1884, 2011.

[19] L. da F. Costa. Comparing cross correlation-based similarities. `https://www.researchgate.net/publication/355546016_Comparing_Cross_Correlation-Based_Similarities`, 2021. [Online; accessed 21-Oct-2021].

[20] L. da F. Costa. Coincidence complex networks. `https://www.researchgate.net/publication/355859189_Coincidence_Complex_Networks`, 2021. [Online; accessed 21-Aug-2021].

[21] L. da F. Costa. Real-valued jaccard and coincidence based hierarchical clustering. `https://www.researchgate.net/publication/355820021_Real-Valued_Jaccard_and_Coincidence_Based_Hierarchical_Clustering`, 2021. [Online; accessed 21-Aug-2021].

[22] L. da F. Costa. Common product neurons. `https://www.researchgate.net/publication/356042155_Common_Product_Neurons`, 2021. [Online; accessed1-Nov-2021].

[23] S. Haykin. *Neural Networks And Learning Machines*. McGraw-Hill Education, 9th edition, 2013.